

Fault Pattern Extraction Via Adjustable Time Segmentation Considering Inflection Points of Sensor Signals for Aircraft Engine Monitoring

Sujeong Baek[†]

Department of Industrial Management Engineering, Hanbat National University

센서 데이터 변곡점에 따른 Time Segmentation 기반 항공기 엔진의 고장 패턴 추출

백수정[†]

한밭대학교 산업경영공학과

As mechatronic systems have various, complex functions and require high performance, automatic fault detection is necessary for secure operation in manufacturing processes. For conducting automatic and real-time fault detection in modern mechatronic systems, multiple sensor signals are collected by internet of things technologies. Since traditional statistical control charts or machine learning approaches show significant results with unified and solid density models under normal operating states but they have limitations with scattered signal models under normal states, many pattern extraction and matching approaches have been paid attention. Signal discretization-based pattern extraction methods are one of popular signal analyses, which reduce the size of the given datasets as much as possible as well as highlight significant and inherent signal behaviors. Since general pattern extraction methods are usually conducted with a fixed size of time segmentation, they can easily cut off significant behaviors, and consequently the performance of the extracted fault patterns will be reduced. In this regard, adjustable time segmentation is proposed to extract much meaningful fault patterns in multiple sensor signals. By considering inflection points of signals, we determine the optimal cut-points of time segments in each sensor signal. In addition, to clarify the inflection points, we apply Savitzky-golay filter to the original datasets. To validate and verify the performance of the proposed segmentation, the dataset collected from an aircraft engine (provided by NASA prognostics center) is used to fault pattern extraction. As a result, the proposed adjustable time segmentation shows better performance in fault pattern extraction.

Keywords : Fault Detection, Time Segmentation, Signal Discretization, Inflection Point, Sensor Data

1. 서 론

생산 시스템들이 자동화됨에 따라 많은 기계장치들에 사물인터넷, 산업용 제어 및 통신 장비들이 함께 부착되

어 메카트로닉스 시스템으로써 기능을 수행하고 있다 [24, 31]. 이러한 메카트로닉스 시스템의 종류가 다양해지고, 수행하는 기능이 복잡해지고 점차 고성능이 필요해짐에 따라 생산 시스템이 안정적으로 생산활동을 수행하기 위해서는 자동화된 고장 감지 및 진단 기술이 매우 중요해졌다[7, 34]. 고장 감지 및 진단과 같은 유지보수는 현장의 전문 작업자에 의해 사후 보전 또는 예방 보

Received 6 August 2021; Finally Revised 13 September 2021;
Accepted 15 September 2021

[†] Corresponding Author : sbaek@hanbat.ac.kr

전의 형태로 많이 이루어지고 있다. 그러나, 최근에는 빅 데이터 분석 및 인공지능 기술이 발달함에 따라 실시간으로 메카트로닉스 장치로부터 센서 데이터를 수집하여 분석하는 상태 기반 모니터링(Condition-based monitoring) 역시 많이 적용되고 있다[6, 12, 14].

기기로부터 수집된 센서 데이터를 분석하여 고장을 감지하는 대표적인 방법으로는 전통적인 통계 기반의 공정 관리 기법 또는 기계학습, 인공지능 기반의 분석 등이 있다. 위와 같은 방법은 기기가 정상적으로 운영되었을 때 수집된 센서 데이터의 양이 충분하고, 정상 상태에 대하여 하나의 확률 모델로 군집화가 가능할 때 고장 감지 모델로서의 효용성을 보이는 편이다[26]. 그러나 실제 기계 시스템으로부터 수집된 센서 데이터를 관찰해보면, 시스템이 정상적으로 운영되고 있음에도 불구하고 해당 정상 신호 모델을 하나의 대표 군집으로 나타내기에는 어려운 경우가 많다[2, 5].

이러한 경우 수집된 센서 신호로부터 정상 상태에 대한 모델이 아닌, 고장 상태와 관련된 유의미한 변동을 추출하는 패턴 추출 및 비교 기반의 고장 감지 기법들이 유의한 감지 성능을 보인다[3, 9]. 센서 데이터 이산화(Signal discretization) 기법은 고장 패턴을 추출하기 위해 사용되는 대표적인 기법으로[3, 16, 33] 시간과 신호 값을 기준으로 입력 데이터의 크기는 줄이면서 신호의 변화는 최대한 유지하는 전처리 기법이다. 대부분의 센서 데이터 이산화 기법에서 신호 값 축의 경우 다양한 방법을 통해 데이터를 분할하나[15, 32], 시간 축의 경우 일반적으로 고정된 크기로 윈도우(Time window)를 사용하여 패턴을 추출한다[10, 32]. 그러나 고정된 크기의 Time segmentation을 수행할 경우 원본 센서 데이터의 변화를 중간에 끊거나, 신호의 변화를 정확하게 표현하지 못하는 경우가 발생하기 쉽다.

그러므로 본 연구에서는 원본 센서 데이터의 변동을 변곡점으로 파악하여 Time segmentation을 수행하는 센서 데이터 이산화 기법을 제안한다. 주어진 신호의 변곡점을 기준으로 Time window를 결정하고 이를 통해 유의미한 고장 패턴을 추출하였다. 제안한 고장 패턴 추출 기법의 성능을 검증하기 위하여 NASA prognostics center에서 제공하는 항공기 엔진 데이터[28, 29]를 활용하였다. 항공기 엔진 데이터에 해당 기법을 적용하여 고장 감지에 유용한 패턴을 추출하고자 한다.

2. 관련 연구

2.1 센서 데이터 기반 고장 감지 연구

기계 설비로부터 수집된 센서 데이터를 이용하여 고

장을 감지하는 기법 중 가장 전통적인 기법은 통계 기반의 공정 관리도를 사용하는 것이다. 예를 들어, Kim et al.[17]은 생산 설비에서 많이 사용되는 구름 베어링의 상태를 실시간으로 감지하기 위하여 통계적 공정 관리 기법을 적용하였다. 그들은 회전하는 구름 베어링에 가속도 센서를 부착하여 센서 데이터를 수집하였으며, 해당 가속도 센서 값을 기반으로 다항 로지스틱 회귀 분석을 적용하였다. 도출한 회귀분석 모델로부터 구름 베어링의 현재 고장 발생 확률을 계산할 수 있었다. 또한, 이산 공정을 위한 작업 중 하나인 프레스 공정에서의 윤활유의 상태를 실시간으로 진단하기 위하여 공정 관리 기법이 적용되었다[21]. 해당 연구에서는 프레스 공정이 원활하게 작동될 수 있도록 윤활유의 상태를 실시간으로 진단하고자, 초음파 센서를 부착하였다. 초음파 센서를 통해 현재 윤활유의 높이를 공정 관리도 기법을 통해 파악하였으며, 해당 수치가 일정 수준보다 떨어질 경우 알람을 발생시켰다. 프로세스 기반 생산 공정에서도 공정 관리 기법을 적용하여 실시간 상태 진단을 적용하였다. 폐수처리 공정에 세 가지 화학특성을 측정하는 센서를 부착한 후 각각의 센서 값에 대해 공정 관리 기법을 적용하였다[18]. 사전에 확인된 정상 상태에서의 히스토리 데이터를 이용하여 공정 관리 모델을 구축하였으며 설정된 신뢰 구간을 벗어날 경우 현재 폐수 처리 공정에 문제가 있다고 판단하였다.

공정 관리 기법을 효과적으로 고장 감지에 적용하기 위해서는 해당 기기가 정상적으로 운영되었을 때의 데이터가 통계적으로 충분하거나, 시스템 운영 상태를 표현하기 위해 필요한 센서 데이터의 수가 상대적으로 적어야 한다[30]. 그러나 현재 많은 생산 공정에서 사용되는 기계 시스템은 하나 또는 소수의 센서 데이터만으로 정상 또는 고장 상태를 파악하는 데는 어려움이 있으며, 통계적으로 신뢰할 만큼의 데이터를 수집하지 못하는 경우도 종종 발생한다[5, 13]. 이러한 경우에는 기계학습 기법을 적용하여 정상 상태에서 수집된 센서 데이터와의 거리 또는 밀도를 기반으로 하여 고장 상태를 분류하는 방법이 많이 적용되고 있다. Ahn et al.[1]은 풍력 터빈의 동력발전기에서 발생하는 고장을 진단하기 위하여 기계학습 기법 중 서포트 벡터 머신(Support Vector Machine, SVM) 기법을 적용하였다. 동력발전기로부터 전류 데이터를 수집하여 위상 값을 계산하였다. 현재의 위상 값을 입력값으로, 시스템의 상태(정상 또는 고장)를 출력값으로 정의한 SVM 모델을 구축하여 고장 감지를 수행하였다. 해당 과정에서 모델 구축을 위해 충분한 데이터를 수집하는 것에 어려움이 있어, K-fold 교차검증 기법을 적용하여 제안하는 고장 감지 모델의 성능을 평가하였다. 또한 기어 펌프의 고장 상태를 감지하기 위해서도 SVM

기법이 적용되었다[20]. 기어 펌프의 고장을 센서 데이터 분석을 통해 감지하기 위해서 먼저 Failure Mode and Effect Analysis(FMEA) 및 Fault Tree Analysis(FTA)를 적용하여 치명적인 고장 상태를 도출하였다. 이 후, 도출된 심각한 고장 상태를 감지하기 위해서 진동 데이터를 수집하였다. 해당 진동 데이터에 푸리에 변환 및 필터 बैं크를 수행하여 특징을 추출하였으며, 최종적으로 SVM에 해당 특징이 입력되면 현재 상태가 정상인지 또는 고장 상태인지 판별할 수 있었다.

SVM이 아닌 다른 기계학습 기법들도 기기의 고장 감지 및 진단을 위해 많이 활용되었다. 예를 들어, Kim et al.[19]은 인버터의 고장을 예측하기 위하여 랜덤 포레스트 기법을 적용하였다. FMEA를 통해 도출된 고장 상태 중 가장 심각도가 높은 과전압을 심각한 고장 상태로 판단하였다. 그 후 수집한 전압 데이터로부터 고차원의 통계적 특징 값을 추출하였다. 추출된 특징 값들을 입력하여 규칙기반의 랜덤 포레스트를 구축하였다. 이를 통해 현재 인버터의 전압 상태가 정상인지 과전압(고장)인지 판단하였다. Lim et al.[22]은 모터의 전류값을 수집하여 전동차 출입문의 상태를 판단하는 연구를 진행하였다. 전동차 자동 출입문에 부착된 모터로부터 전류값을 실시간으로 수집하였으며, 문이 열릴 때와 닫힐 때를 구분하여 주파수 파형을 추출하였다. 각각의 열림, 닫힘 상태에서 유의한 통계 특성치를 계산하여, 고장 발생 유무를 판단하기 위한 K-Nearest Neighborhood 알고리즘의 입력값으로 사용하였다. 해당 고장 감지 모델을 이용하여, 평균 전류와 전류 값의 첨도 치가 높다면 모터의 권선이 단락되어 과부하가 발생함을 발견할 수 있었다.

이러한 기계학습 기반의 방법 역시 기기가 정상적으로 운영되었을 때의 센서 데이터가 상대적으로 하나의 확률 모델로 표현할 수 있을 때 고장 감지 모델로서의 효용성을 보이는 편이다[26]. 그러나 실제 메카트로닉스 시스템으로부터 수집된 센서 데이터의 경우, 시스템이 정상적으로 운영되고 있음에도 불구하고 해당 정상 신호 모델을 하나의 군집으로 나타내기에는 어려운 경우가 많다[2, 5]. 그러므로, 이러한 경우에는 정상 상태에서 대표 모델을 구축하는 것이 아니라 시스템에서 고장이 발생했을 때에 대한 모델을 구축하는 것이 바람직하다.

패턴 추출 기반의 고장 감지 기법은 시스템에서 고장이 발생하였을 때의 센서 데이터를 정상 상태에서 수집된 값과 비교하여 고장 상태에서만 나타나는 유의미한 변동을 추출하는 방법이다. 해당 방법은 정상 상태에서의 데이터가 굉장히 산포되어 있거나, 특정 소수의 센서의 변화가 시스템의 상태 변화와 유의미한 관계를 갖지 않을 때, 또는 많은 센서 데이터에 대해 분석을 동시에 수행하여야 할 때 유의한 감지 성능을 보인다[3, 9]. 다시

말해, 패턴 기반의 고장 감지 기법의 경우, 고장 상태에서만 발생하면서 동시에 정상 상태에서는 발생하지 않는 센서 신호의 특성을 추출해내기 때문에 상대적으로 기계학습 기법에 비해 False alarm을 내는 측면에서 강건한 효과를 얻을 수 있다. 또한 기계학습 및 인공지능 기반의 분석 기법의 경우, 실제 수집된 신호 값의 거리에 기반하여 고장을 감지하게 된다. 이 경우 새롭게 수집된 데이터가 기존 학습된 모델 중 거리가 가까운 상태 값으로 분류가 된다. 그러나 복잡한 기계 시스템일수록 데이터가 굉장히 산포되어 있어 이러한 거리 값을 통한 유사성을 비교할 때 정확도가 낮아질 수 있다. 위와 같은 문제 역시 패턴 기반의 기법을 사용할 경우 기존 데이터와 신규 데이터 간의 비교를 거리 값을 사용하지 않기에[25], 데이터 양상을 관찰하였을 때 산포의 정도가 크더라도 고장 감지의 높은 정확도를 기대할 수 있다는 장점이 있다.

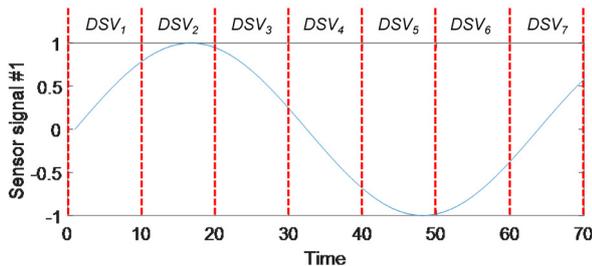
예를 들어 Bai et al.[4]의 경우, 데이터 이산화 기법 중 평균값을 토대로 원본 센서 데이터를 변환하는 Symbolic Aggregation Approximation 기법을 온도 센서 채널에 적용하였다. 그 후, 변환된 라벨을 정해진 시간 축 크기로 결합하여 단어를 생성하였다. 생성된 단어 중 고장 상태에서만 통계적으로 유의한 수준으로 많이 발견되는 단어를 고장 패턴으로 정의하여 고장 감지에 적용하였다. Duan et al.[8]의 연구에서도 시스템의 정상 상태와 스프링의 고장 상태, 밸브의 고장 상태를 구별하기 위해서 고장 패턴을 추출하였다. 데이터 이산화 기법을 적용한 후, 추출된 라벨들이 발생한 횟수를 기반으로 Heat map을 생성하였다. 해당 Heat map 중 가능한 고장 상태에서만 발생하는 Heat map을 고장 패턴으로 정의하여 고장 감지를 위해 사용하였다.

2.2 데이터 이산화 기법에서의 Time Segmentation

고장 패턴 추출을 위해 적용되는 데이터 이산화 기법에서 시간 축에 수행되는 과정을 Time segmentation이라 부른다. 해당 과정은 연속된 시간 축 상에 존재하는 수집된 원본 센서 데이터를 이산화하는 것을 의미한다. 다시 말해, n 길이의 원본 시계열 데이터의 시간 축을 주어진 기준에 따라 분할하여, s 개의 Time segment로 변환하는 과정이다($n \gg s$). 위와 같은 관점에서 데이터를 수집하여 전처리 과정에서 수행하는 Down-sampling과 흡사하다고 할 수 있다.

가장 대표적으로 많이 사용하는 Time segmentation 기법은 고정 길이를 이용하여 시간 축을 분할하는 기법이다(Time segmentation with the fixed length)[10]. 이는 사전에 정의된 Time window의 길이(w)를 이용하여 분할되는 인접한 Time segment가 서로 겹치지 않게 시간 축을

나눈다. 그러나 고정된 크기로 Time segmentation을 진행하게 되면, 신호가 동일한 Trend를 보이고 있음에도 이를 중간에 끊어 다른 DSV로 분할하거나, 두 가지 이상의 신호 변화를 하나의 DSV로 합치는 등 센서 신호의 변화를 효과적으로 표현하지 못하는 경우가 발생할 수 있다. 예를 들어, <Figure 1>의 경우 간단한 사인과 형태의 신호가 현재 원본 데이터로 수집되고 있다. 이 때 10개의 센서 값을 하나의 Time window로 정의한다면, DSV_2 와 같이 센서 데이터가 증가하다가 감소하는 경향을 하나의 DSV로 합쳐서 라벨로 변환하는 과정에서 신호 기울기 정보가 없어지기도 한다. 또한 DSV_3 , DSV_4 와 같이 하나의 DSV가 아닌 두개의 DSV로 정의됨으로써 패턴 추출에서의 시간적, 감지 성능적 효율성이 감소하는 경우도 발생한다.



<Figure 1> Time Segmentation with a Fixed Size of Time Windows (= 10 Time Points)

위와 같은 단점을 해결하기 위해서 많은 연구들이 진행되어 왔다. 그 중 대표적인 방법이 동적(Dynamic) Time segmentation 기법이다[15]. 해당 기법은 시간 축 분할을 Bottom-up 방식으로 수행하여 크기가 각각 다른 Time segment를 획득하는 것이다. 예를 들어, 가장 작은 크기의 Time window를 설정하여 모든 센서 데이터를 고정 크기로, 겹치지 않게 분할한 후, 인접한 Time window의 대표 값(예: 평균값, 중앙값 등)이 동일하다면 하나의 Time window로 합치는 것이다[10]. 그러나 여전히 가장 작은 Time window의 크기에 따라 센서 데이터의 변동이 적절하지 않게 분할될 확률이 높고, 이를 해결하기 위해 반대로 매우 작은 Time window 크기를 사용하게 될 경우 데이터 이산화 기법을 적용하는 효율성(예: 데이터의 양 축소)이 감소하게 된다.

Sliding window의 경우에도 센서 데이터의 변동을 최대한 보존하고자 적용되는 기법 중에 하나이다. 고정된 Time window 길이를 이용하나 인접한 Time segment간 일정한 양의 데이터 포인트를 공유하게끔 하는 방법이다. 이렇게 될 경우 기존의 고정 길이 분할법에 비해서는 센서 데이터의 변동을 많이 유지할 수 있고, 신호의 잡음

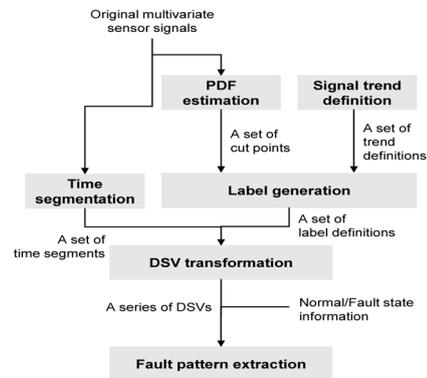
이 많은 경우에는 좋은 성능을 보이는 편이나, 전처리된 센서 데이터를 사용하거나, 센서 데이터의 양이 다소 방대한 경우에는 유의한 효과를 보기가 어렵다 [15].

그러므로 본 연구에서는 센서 데이터의 변화를 최대한 해치지 않으면서도, 데이터 이산화의 효과를 유지할 수 있는 Adjustable time segmentation 기법을 제안하고자 한다. 자세한 내용은 다음 장에서 이어서 설명하겠다.

3. 센서 데이터의 변곡점을 고려한 고장 패턴 추출 기법

3.1 센서 데이터 이산화 기법을 통한 고장 패턴 추출

수집된 다변량 센서 데이터로부터 고장 패턴을 추출하기 위해서 Baek and Kim[3]이 제안한 데이터 이산화 기법을 적용하였다. 해당 고장 패턴 추출 기법은 <Figure 2>에서 보여지는 바와 같이 수행되며 상세한 설명은 다음과 같다.



<Figure 2> The Procedure of Fault Pattern Extraction Via Multivariate Data Discretization(DSV Indicates a Discretized State Vector)

메카트로닉스 시스템으로부터 수집된 원본 다변량 데이터를 이용하여 어떻게 Discretized State Vector(DSV)를 정의할 것인지를 결정하기 위하여 가장 먼저 확률 밀도 함수를 추정한다(Probability density function(PDF) estimation). 기존 통계적 공정 관리도 등의 방법을 이용하여 수집된 센서 데이터를 정상 상태와 고장 상태로 분류하는 경우, 원본 센서 데이터들의 확률 밀도 함수를 일반적으로 정규 분포로 가정한다. 그러나 메카트로닉스 시스템에서 수집된 센서 데이터는 서론에서 언급된 바와 같이 정상 상태와 고장 상태가 모두 하나의 군집으로 표현되기 어렵

기 때문에, 굉장히 산포된 형태인 Sparse한 밀도 함수를 보이는 경우가 많다. 그러므로 각각의 원본 센서 신호에 대해 최적의 확률 밀도 함수를 추정한 후, 센서 축의 구간의 수(The number of bins) 및 가장 큰 구간의 비중(Bin width threshold)에 따라 Bin을 결정하기 위한 구간 경계점들(A set of cut-points)을 계산한다. Cut-point는 전체 센서 데이터를 정해진 Bin(센서 축에서 분할되는 구간의 수)으로 구별할 때 사용하는 기준을 의미한다. 예를 들어 평균값을 이용하여 센서 데이터 축을 분할한다면, 한 개의 Bin에 대한 최소값, 최대값이 Cut-point가 될 것이다. 이를 이용해 주어진 센서 신호의 평균값이 어떤 센서 구간에 속하는지 확인하여 라벨(Label)을 부여하게 되는 것이다. 라벨은 원본 센서 데이터가 DSV가 되기 위해 변환되는 과정에서 원래의 센서 값이 속하는 Bin이 어떤 것인지 알기 위한 대표 값이다. 예를 들어 평균값을 이용하여 센서 축을 나누었고, 현재 Time segment의 원본 데이터가 2번째 위치한 Bin에 속한다면, 라벨 값은 2 또는 l_2 등으로 나타낼 수 있다.

DSV로 원본 데이터를 변환하기 위해 라벨을 생성할 때, 센서 데이터의 평균 크기뿐만 아니라, 신호의 변화를 사용할 수도 있다. 1차 기울기, 2차 기울기 등을 고려하여 신호의 시간에 따른 변화(Signal trend)를 라벨 정의 시 포함할 수도 있다. 이외에도 신호의 엔트로피[23], 카이 제곱 값(Chi2)[11] 등을 추가로 더 고려하여 라벨을 정의할 수도 있다. 라벨 정의를 위한 신호의 특징 값(Feature)이 결정되고 나면 해당 Feature들을 통해 얻을 수 있는 모든 조합을 도출하여 DSV 생성을 위한 라벨을 정의한다(Label generation).

라벨을 정의하면서, 시간 축을 나누기 위한 Time segmentation을 동시에 수행하여야 한다. 전통적으로는 고정된 크기를 사용하여, 각각의 Time segment가 겹치지 않게끔 신호를 분할한다. 이 후 각각의 Time segment 내에 포함된 센서 신호와 사전에 정의된 라벨의 정의와 비교하여 원본 데이터를 라벨 값으로 변환한다(DSV transformation). 다시 말해 하나의 Time segment 내의 센서 신호가 행의 수는 센서의 수이며 열은 하나인 벡터 형태로 변환되는 것이다.

$$X = [X_1, X_2, X_3, \dots, X_m]$$

$$X_i = [x_{1i}, x_{2i}, x_{3i}, \dots, x_{mi}]$$

$$DSV(X) = [dsv_1, dsv_2, dsv_3, \dots, dsv_s]$$

where m is the number of sensors, n is the length of a time-series signal, s is the number of time segments ($s = \left\lceil \frac{n}{w} \right\rceil$), w is the length of a time segment(window)

모든 센서 데이터를 DSV로 변환하고 나면, 해당 DSV들이 발생하였던 시점의 기계 시스템의 상태정보(Normal/Fault state information)에 따라 고장 패턴(Fault pattern)을 다음과 같이 정의한다[3]:

‘시스템의 정상 상태에서는 발견되지 않고, 고장 상태에서에서만 발견되는 DSV.’

예를 들어 다음과 같다. 원본 센서 데이터 X로부터 다변량 데이터 이산화를 적용할 결과 총 10개의 DSV 벡터를 획득하였다고 가정해보자. 이때 고장이 발생하였을 때 수집된 센서 데이터에서 변환된 DSV가 4가지이고, 시스템이 정상 상태일 때 수집된 센서 데이터로부터 획득한 DSV가 다음과 같은 8가지라면 정상 상태에서는 발견되지 않으나 고장 상태에서에서만 발견되는 2개의 DSV가 최종적으로 고장 패턴으로 추출되는 것이다.

$$DSV = \{dsv_1, dsv_2, dsv_3, \dots, dsv_{10}\}$$

$$DSV_{faulty} = \{dsv_1, dsv_2, dsv_3, dsv_4\}$$

$$DSV_{normal} = \{dsv_3, dsv_4, dsv_5, dsv_6, dsv_7, \\ dsv_8, dsv_9, dsv_{10}\}$$

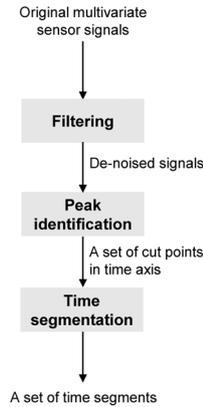
$$FP = \{dsv_1, dsv_2\}$$

where FP is the set of the extracted fault patterns

이렇게 정의된 고장 패턴은 시스템의 고장이 발생했을 때만 관찰되는 짧은 Time window 내의 센서 신호의 조합이다. 그러므로 메카트로닉스 시스템 운영 중 시간으로 센서 데이터를 수집하면, 고장이 발생하였을 때 즉시 감지할 수 있는 지표로 사용할 수 있다. 또한 현재는 전통적인 집합이론(Crisp set)을 통해 고장 패턴을 추출하였으나, 전문가의 의견을 고려하여 해당 과정을 적절한 기계학습이나 인공지능 등을 이용한 분류기로 대체할 수 있다[25].

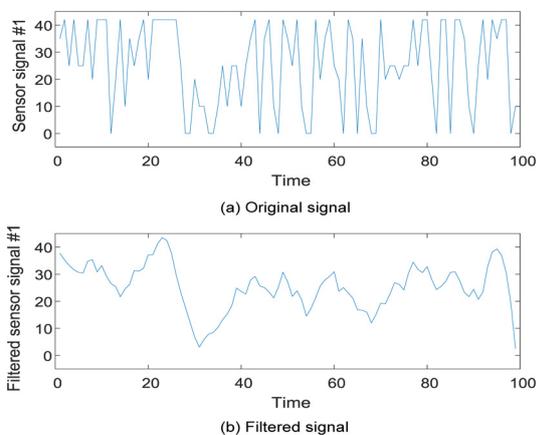
3.2 센서 데이터변곡점을 고려한 Adjustable Time Segmentation

제3.1절에서 설명한 기존의 고장 패턴 추출 기법에서는 고정된 크기의 Time window를 사용하여 원본 데이터를 DSV로 변환한다. 그러나 제2.2절에서 언급한 바와 같이 고정된 크기로 Time segmentation을 진행하게 되면, 내제된 신호의 변화를 왜곡한 DSV로 추출될 확률이 매우 높다. 그러므로 본 연구에서는 센서 데이터의 변곡점을 파악하여 이를 토대로 DSV로 변환하기 위한 Adjustable time segmentation을 제안하고자 한다. 제안하는 Adjustable time segmentation 기법은 <Figure 3>과 같다.



<Figure 3> The Procedure of the Proposed Adjustable Time Segmentation

가장 먼저 수집된 원본 다변량 센서 데이터에 잡음(Noise)을 제거하기 위하여 필터링(Filtering)을 수행한다. 신호의 잡음은 대상으로 하는 메카트로닉스 시스템의 작동 상태의 변화로 인해 발생한 변화가 아닌, 수집 기기에서 발생한 의도치 않은 특이치, 누락 값, 중복 값 등과 같이 신호 분석에 유의미한 영향을 주지 않는 것들을 의미한다. 이러한 잡음은 신호 필터를 이용하면 제거할 수 있다. 본 연구에서는 사비츠키-골레이(Savitzky-Golay) 필터를 적용하였다. 대부분의 표준 평균 Finite impulse response(FIR) 필터의 경우 높은 주파수대의 변화를 신호의 잡음으로 판단하여 이를 제거하는 경향이 있다[27]. 그러나 사비츠키-골레이 필터의 경우 <Figure 4>와 같이 필터링된 출력 신호와 원본 입력 신호 간의 최소 제곱 오차를 최소화하면서도 고주파수 대역의 신호를 잡음과 분리하여 유지한다[27]는 면에서 센서 데이터의 유의미한 변동점을 파악하는 데 적절한 것으로 판단되어 본 연구에서 적용되었다.



<Figure 4> The Original and the Filtered (by Savitzky-Golay filter) Sensor Signal to not Only Identify Peak Points But Also Remove Inherent Noise

이 후 필터링된 신호에서 인접한 신호 값들의 비교하여 변곡점을 찾아, 이를 Time segmentation을 위한 구별점으로 사용하였다. 전후의 인접한 신호의 변화를 확인하여 해당 값이 극대값인지, 극소값인지 확인하는 일반적인 신호 처리 과정을 통해 변곡점을 찾았다. 해당 극점을 각각의 센서 데이터에 대해 적용한 후, 동일한 시간대의 모든 센서에 대해 동일하게 적용하여 DSV로 최종적으로 변환하였다. 이후 기존의 고장 패턴 추출 기법에 따라 고장 패턴을 최종적으로 도출하였다. 다시 말해, 제안하는 Adjustable time segmentation 과정은 <Table 1>에 요약된 Pseudo code와 같이 수행하였다.

<Table 1> The Pseudo Code for the Proposed Adjustable Time Segmentation

Variables	m = the number of sensors n = the length of time series signals X_i = i^{th} sensor signal Y = a set of time segments
1	for $i = 1 : m$:
2	Filtered _{i} ← savitzky_golay_Filtering(X_i)
3	peaks _{i} ← find_Peaks(Filtered _{i})
4	end for
5	Peaks ← [Peaks; n]
6	Final_peaks ← find_Unique_Value(Peaks)
7	Y ← time_Segmentation(X , Final_peaks)

4. 항공기 엔진의 고장 패턴 추출 결과

4.1 항공기 엔진 데이터

제안한 Adjustable time segmentation 기법의 효용성을 평가하고자 NASA prognostics center에서 제공하는 항공기 엔진 데이터[28, 29]를 분석하였다. 사용한 데이터는 항공기 엔진의 열화를 판단하기 위해서 Commercial Modular Aero-propulsion System Simulation(C-MAPSS)을 이용해 개발된 항공기 터보팬 엔진 시뮬레이터에서 수집한 센서 신호로 구성되어 있다. 아래의 <Table 2>와 같이 원본 데이터는 하나의 엔진 운영 주기마다 엔진 번호, 작동 시간을 포함하여 총 26개의 데이터를 샘플링 주기에 맞춰 수집되고 있다. <Table 2>의 Operation setting value는 시뮬레이터의 입력 파라미터 값으로 사전에 미리 정의된 시스템의 특성 값이다. Analog sensor measurement는 엔진이 작동하는 동안 시간에 따라 상태가 변하는 센서 데이터를 의미한다. 또한 엔진 가동부터 엔진이 더 이상 작동할 수 없는 상태까지 센서 데이터를 수집하였다. 해당 센서 데이터 중 엔진 번호, 작동 시간 및 시간에 따른 변동이 전혀 없는 센서 데이터(수

집된 센서 신호의 분산이 0인 경우)는 분석에서 제외하여 최종적으로 본 연구에서는 17개의 아날로그 센서 신호만을 고장 패턴 추출을 위해 사용하였다. Nasa 사에서 제공하는 데이터 설명서에 따르면, 수집된 센서 데이터는 팬 입구의 온도, 고압 컴프레서(HPC Compressor)의 온도, 팬 입구의 압력, 팬의 실제 회전 속도, 팬의 이론상 회전 속도 등의 값을 의미하나, 안타깝게도 경진대회를 위해 데이터를 생성하였기 때문에, 실제 어떤 Variable ID가 어떤 센서에 해당되는지에 대해서는 공개되어 있지 않다[29].

<Table 2> Description of the Collected Sensor Signals from an Engine Simulator

Variable ID	Description
1	Engine ID
2	Operating time
3	Operation setting value 1
4	Operation setting value 2
5	Operation setting value 3
6	Analog sensor measurement 1
...	...
26	Analog sensor measurement 21

항공기 엔진의 고장 상태를 재현하기 위하여 원본 데이터 생성자는 임의적으로 고압 컴프레서의 고장만을 발생시키거나(1개의 고장 모드), 고압 컴프레서 고장과 팬의 열화(Fan degradation)를 함께 발생시켰다(2개의 고장 모드). 또한 데이터 세트는 크게 운영 조건에 따라 두 가지 종류로도 구별할 수 있다. 운영 조건은 주어진 센서 신호 중 Operation setting value에 따라 결정된다. 본 논문에서는 원본 데이터를 제공하는 Nasa 사의 설명서에서 기재된 바 같이 두 가지 운영 조건을 Condition 1, Condition 6으로 구별하고자 한다[29]. Condition 1은 Sea level에 해당되며 고도가 0 ft, 마하 값이 0 mac, 그리고 Sea-level 온도가 100°F인 경우를 의미한다. 반면에 Condition 6은 운행 중인 경우로 고도가 0 ft에서 최대 42ft까지 임의적으로 변경되었으며(평균 24ft), 마하 값 역시 0에서 0.84사이 값을 띄었다(평균 0.57). Sea-level 온도 역시 60°F에서 100°F 사이에서 계속 변경되었다(평균 94°F). 요약하자면 총 4개의 데이터 세트를 본 연구에서 제안된 기법을 검증하기 위해 사용하였으며 <Table 3>과 같다. 하나의 데이터 셋에는 고장 패턴 추출을 위한 여러 개의 시계열 데이터가 포함되어 있으며 이를 이용해 데이터셋 별로 고장 패턴을 추출하였다.

고장 패턴 추출을 위하여 시스템의 상태는 정상과 고장으로 분류하였으며 이는 원본 데이터에서 주어지는 잔여 유효 수명을 기반으로 수행하였다. 26개의 값이 일정

한 주기마다 수집되어 저장된 하나의 시계열 데이터는 엔진이 정상적으로 작동했을 때부터 고장으로 인해 더 이상 작동할 수 없을 때까지 수집되었으므로, 이를 역추적하여 각 Time stamp 별로 잔여 유효 수명을 계산한다. 하나의 Time stamp를 하나의 Cycle로 간주하여, ‘예상치 못한 비정상적 정지가 발생하기 50 cycle전부터 실제 고장 상태가 나타났다.’라고 정의하였다. 예를 들어, <Figure 5(b)>에서 Analog sensor #2의 경우 엔진의 초기 상태부터 더 이상 작동이 하지 않을 때까지 운행한 결과 총 약 150 cycle 동안 운행한 것을 확인할 수 있다. 이 중 100 cycle 제부터는 고장 상태로 정의하였다(비정상적인 정지까지 남은 잔여 유효 수명이 50 cycle 이하). 반대로 처음 시작부터 99 cycle까지는 남은 잔여 유효 수명이 50 cycle보다 크므로 정상 상태로 정의하였다.

<Table 3> The Used Data sets for Validating the Performance of the Proposed Approach

Dataset ID	Operating condition	Fault modes	The number of time-series data
Dataset 1	Condition 1	HPC degradation (one fault mode)	100
Dataset 2	Condition 1	HPC and fan degradation (two fault modes)	100
Dataset 3	Condition 6	HPC degradation (one fault mode)	260
Dataset 4	Condition 6	HPC and fan degradation (two fault modes)	248

4.2 Adjustable Time Segmentation을 통한 고장 패턴 추출 결과

제안한 Adjustable time segmentation 기법을 적용한 고장 패턴 추출 기법의 성능을 확인하기 위하여 기존의 추출 기법과 비교를 하고자 한다. 고장 패턴의 성능은 총 고장 상태의 수 대비 고장 패턴으로 감지한 고장 상태의 수의 비율로 나타낼 수 있다 [3].

$$\frac{\text{The number of discernible fault state} = \text{the number of fault states where any fault pattern is observed}}{\text{the number of the given fault states}}$$

그러므로 해당 연구에서는 본 Discernible fault state 의 값으로 고장 패턴 추출법의 성능을 평가하고자 한다. 다시 말해 해당 지표는 실제 고장 상태 중 고장 패턴을 통해 몇 개의 고장 상태를 정확히 감지하였는지를 의미하며, True positive rate에 해당된다. 고장 감지에 있어서는

False positive rate에 해당되는 정상 상태에서 고장 패턴이 발견되어 고장 상태로 잘못 감지하는 것 역시 중요한 척도이다. 그러나 본 연구에서 적용된 고장 패턴 추출 기법에 의하면 (추출된 고장 패턴 정의의 상) 고장 패턴이 정상 상태에서 등장할 수는 없으므로 이에 대한 계산은 생략하였다.

제안한 기법과 기존의 고장 패턴 추출 기법에서 사용자가 공통적으로 정의해야 하는 파라미터는 다음과 같다:

- The number of bins: 하나의 센서 축에서의 구간의 수
- Bin width threshold: 중앙 구간이 전체 확률밀도함수에서 차지하는 비율
- LinearT: 선형 기울기의 라벨 포함 여부

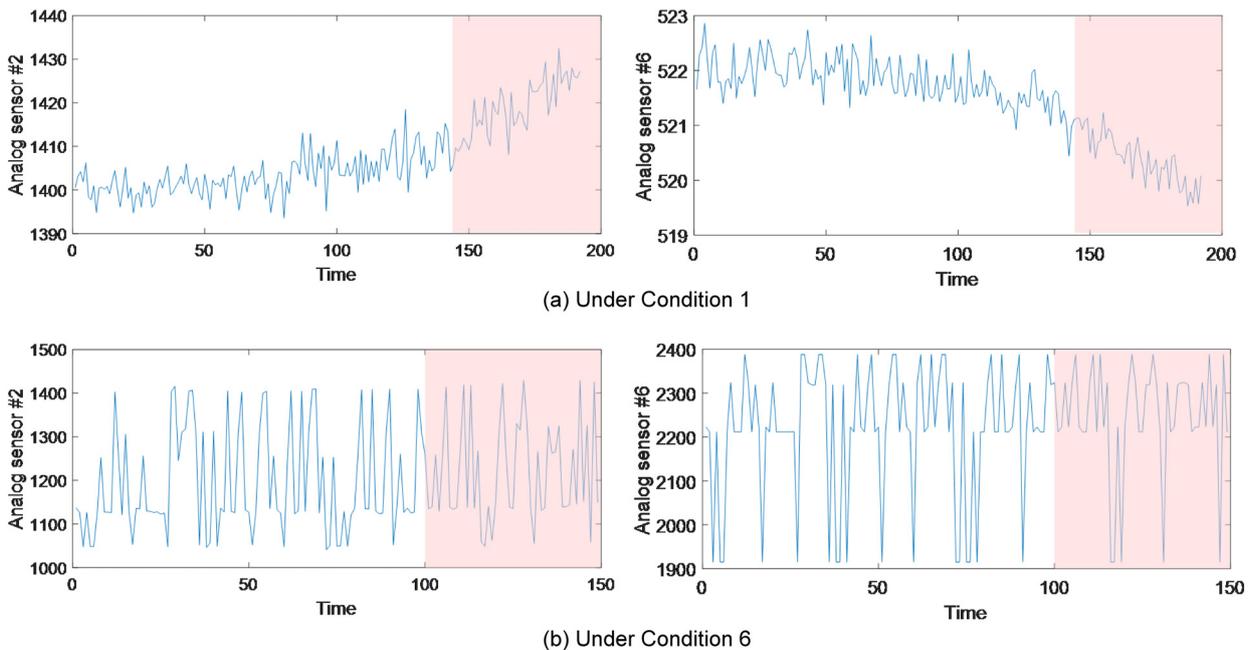
해당 파라미터들의 경우 사전 연구에서 파라미터 값에 따른 고장 성능의 변화에 대해 민감도 분석을 수행하였으며, 이를 통해 최적 값을 도출하였다[3]. 본 연구에서도 센서 축 분할의 경우 동일한 과정을 거쳤으므로 해당 연구 결과를 통해 얻은 최적 파라미터 값을 기존과 제안한 기법에 동일하게 적용하였다. 예를 들어 각각의 센서 축은 7개의 Bin으로 구별하였으며, Bin width threshold는 50%로, 그리고 LinearT의 경우 가능한 다양한 고장 패턴을 추출하기 위해 포함하였다(True). Adjustable time segmentation 기법의 경우 매우 짧은 Time window가 도출될 경우 패턴 추출에 있어 시간 소모량이 매우 증가하게 되므로, 극대점만을 필터링된 센서

신호에서 계산하여 Time segmentation을 수행하였다.

각 항공기 엔진 데이터 셋에 따른 고장 패턴 추출 성능은 <Table 4>와 같이 도출되었다. 먼저, Dataset 1과 Dataset 2에 해당되는 Condition 1의 상황에서는 기존의 기법과 제안된 Adjustable time segmentation을 적용했을 경우 모든 고장 상태에 대해 감지 가능한 고장 패턴을 추출할 수 있었다. 이는 Condition 1의 경우 <Figure 5(a)>에서 보여지는 바와 같이 다소 정상 상태에서 고장 상태로 변화하는 열화 현상을 특정 원본 신호에서 유추가 가능한 편이기 때문인 것으로 판단된다. 다시 말해 Time segment 내에서의 시간에 따른 센서 신호의 변화를 유의미하게 고려하지 않더라도 센서 신호 값의 평균 크기의 차이로 유의미한 고장 패턴을 추출할 수 있는 경우라 할 수 있다.

<Table 4> The Percentage of Discernible Fault States by the Existing and Proposed Time Segmentation Approaches (The number in parentheses is the exact number of discernible fault states over the entire number of fault states)

	The existing fixed size of time segmentation	The proposed adjustable time segmentation
Dataset 1	100% (100/100)	100% (100/100)
Dataset 2	100% (100/100)	100% (100/100)
Dataset 3	63% (163/260)	90% (233/260)
Dataset 4	70% (173/248)	95% (236/248)

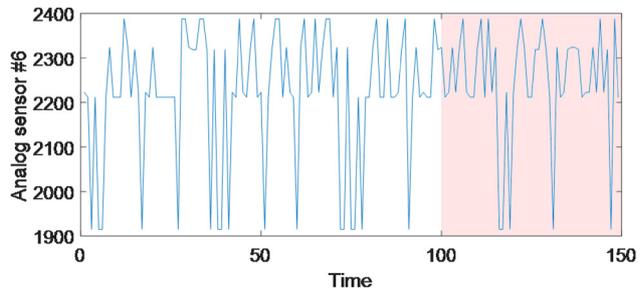
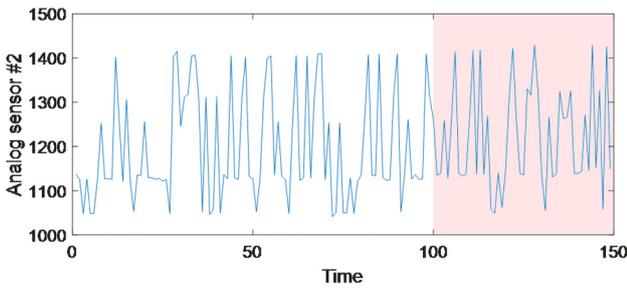


<Figure 5> Original signal of sensor #2 and #6 in Dataset 3, and the corresponding labels transformed by the existing and proposed time segmentation (normal state filled with white, whereas fault state filled with red)

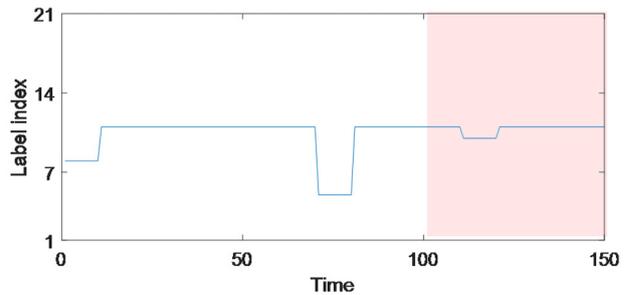
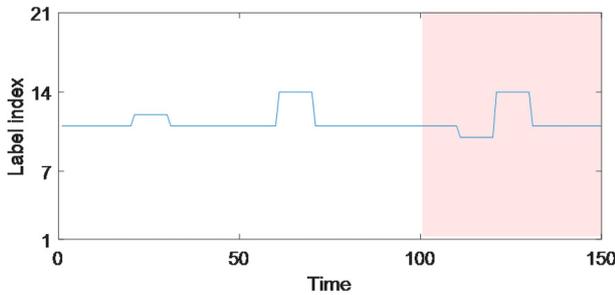
반면, Dataset 3과 Dataset 4에 해당되는 Condition 6에서는 제안된 기법을 사용하였을 때 주어진 고장 상태를 더 많이 감지할 수 있음을 확인할 수 있다. 해당 운영 조건은 열화가 발생함에 따라 센서 데이터에서도 상승 또는 하강과 같은 뚜렷한 센서 신호의 변동이 확인되었던 Condition 1과 달리 시스템에 상태에 따라 특이한 센서 변동이 보이지 않았다. 이는 <Figure 5(b)>을 통해 확인할 수 있다. 이와 같은 센서 데이터의 특성으로 인하여 기존의 고정 길이의 Time segmentation을 적용하였을 때 <Figure 6(b)>와 같이 정상과 고장상태에서의 센서 데이터의 변동을 구별하기 어려운 DSV로 변환되게 된다. 그러나, 제안된 Adjustable time segmentation 기법을 사용한 경우, <Figure 6(c)>와 같이 추출되는 DSV도 다양해지고, 정상과 고장 상태에서 나타나는 DSV 간에 차이가 발생한다. 다시 말해, Time window 내의 시간에

따른 센서 신호의 변화를 훼손시키지 않았을 때 고장 패턴의 성능이 유의미하게 상승하였던 것으로 판단된다. 이외에도 시스템 열화를 위해 발생시킨 고장 모드의 수와 종류의 경우에 따른 제안된 기법의 성능 차이는 유의하게 발생하지 않았다.

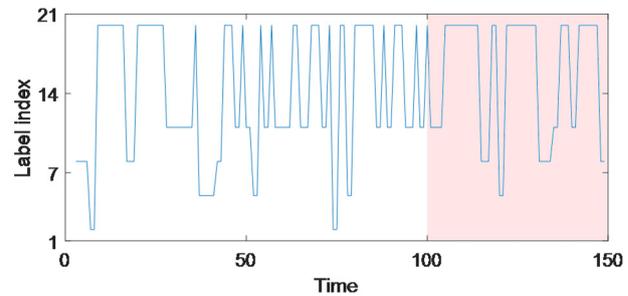
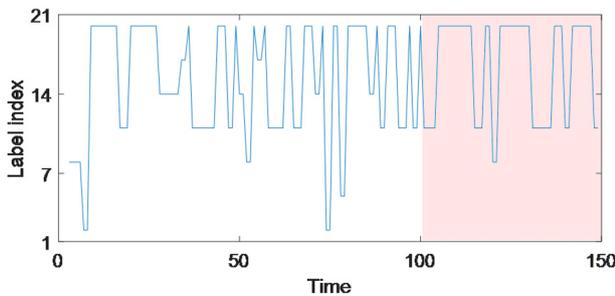
또한 추가적으로 추출된 고장 패턴 중 시스템의 고장 상태 발생과 가장 큰 영향을 미치는 센서 데이터를 밝혀 내기 위하여 DSV 내에서 센서 축 별 라벨의 변화를 관찰하고자 하였다. 그러나 시스템의 정상상태에 비해 고장 상태에서 더 많은 변동을 보이는 센서, 또는 고장 상태에서에서만 특별한 라벨 값을 보이는 센서는 발견할 수 없었다. 이는 하나의 센서 값 변화가 고장 상태를 발생시키는 것이 아니라 여러 센서 값의 변화의 조합이 고장 상태의 발생과 연관되어 있음을 의미한다.



(a) Original signal



(b) The obtained labels in a sensor signal by the fixed time segmentation



(c) The obtained labels in a sensor signal by the adjustable time segmentation

<Figure 6> Analog signal of sensor #2 and #6, and the corresponding labels for DSV generation by the existing and proposed time segmentation collected as Dataset 3 (normal state filled with white, whereas fault state filled with red)

5. 결 론

본 연구에서는 메카트로닉스 시스템의 고장 상태를 감지하고자 수집된 아날로그 센서 신호 데이터를 통해 고장 패턴을 추출하였다. 해당 과정에서 시간에 따른 센서 신호의 변동을 최대한 패턴 내에 유지하고자 기존과 달리 Time window의 크기를 상황에 맞춰 변경하였다. 원본 센서 신호로부터 변곡점을 확인하여 이를 통해 각 DSV의 Time window 크기를 최적화하였다. 그 결과 항공기 엔진에서 발생한 고압 컴프레서 또는 팬 고장에 대해 기존의 고정된 크기의 Time window를 통한 고장 패턴을 추출하였을 때 보다 본 연구에서 제안한 Adjustable time segmentation 기반의 고장 패턴 추출법을 적용하였을 때 더 좋은 고장 감지 성능을 보임을 검증하였다.

그러나 해당 Time window 길이 변화를 실시간 모니터링에 적용하기 위하여 변곡점 탐색 기법 및 다른 길이의 DSV간 매칭에 관한 추가 연구가 계속해서 필요하다. 예를 들어 DSV간의 Time window 길이가 모두 다르므로, 다양한 크기의 고장 패턴과 실시간으로 수집된 DSV가 동일한지 판단하기 위한 유사도 평가 방법에 대한 연구가 추가적으로 필요할 것이다. 현재 해당 연구는 하나의 항공기 엔진에서 발생하는 고장 상태에 대해 고장 패턴을 추출하였으나, 이와 흡사한 기계시스템에서 수집된 데이터를 분석하였을 때도 유의미한 고장 패턴을 추출할 수 있는 지에 대하여 검증이 수행될 필요가 있다. 또한 고장 패턴을 추출하기 전 입력되는 센서 데이터 중 어떤 센서 데이터가 고장 상태를 감지하는데 가장 큰 영향을 미치는 지에 대한 전처리 연구는 많이 진행되고 있다 [2]. 그러나, 추출된 고장 패턴에서도 정상 상태에서 발생하는 패턴 등과 비교하여 어떤 센서가 가장 큰 변화를 보이고 있는지에 대하여 본 연구에서 수행한 통계분석 수준이 아닌 데이터 마이닝 기반의 분석이 이루어졌을 때 고장 원인을 발견하고 조치를 취하는 고장 진단에 더 큰 도움을 줄 수 있을 것으로 예상된다.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2019R1G1A1097478), This work was also supported by project for Industry-Academic Cooperation Based Platform R&D funded Korea Ministry of SMEs and Startups in 2020. (Project No. S3025825)

References

- [1] Ahn, B., Kim, B., Jo, H., and Kim, I., A Study on the Fault Diagnosis of Wind Turbines Using SVM Based on Cross Validation, *Proceedings of 2021 Spring Conference of The Korean Institute of Electrical Engineers*, 2021, JeJu, Republic of Korea, pp. 136-137.
- [2] Baek, S. and Kim, D.-Y., Abrupt Variance and Discernibility Analyses of Multi-Sensor Signals for Fault Pattern Extraction, *Computers & Industrial Engineering*, 2019, Vol. 128, pp. 999-1007.
- [3] Baek, S. and Kim, D.-Y., Empirical Sensitivity Analysis of Discretization Parameters for Fault Pattern Extraction from Multivariate Time Series Data, *IEEE Transactions on Cybernetics*, 2017, Vol. 47, No. 5, pp. 1198-1209.
- [4] Bai, X., Xiong, Y., Zhu, Y., Liu, Q., and Chen, Z., Co-anomaly Event Detection in Multiple Temperature Series, *Proceedings of the 6th International Conference of Knowledge Science, Engineering, and Management*, 2013, pp. 1-14.
- [5] Chandola, V., Banerjee, A., and Kumar, V., Anomaly Detection for Discrete Sequences: A Survey, *IEEE Transactions on Knowledge and Data Engineering*, 2012, Vol. 24, No. 5, pp. 823-839.
- [6] Cheon, K.M. and Yang, J., An Ensemble Model for Machine Failure Prediction, *J. Soc. Korea Ind. Syst. Eng.*, 2020, Vol. 43, No. 1, pp. 123-131.
- [7] Cho, S.-J., Jun, H.-B., Shin, J.-H., and Hwang, H.-J., A Study on Estimating the Next Failure Time of a Compressor in LNG FPSO, *J. Soc. Korea Ind. Syst. Eng.*, 2014, Vol. 37, No.4, pp. 12-23.
- [8] Duan, L., Zhang, Y., Zhao, J., Wang, J., Wang, X., and Zho, F., A Hybrid Approach of Sax and Bitmap for Machinery Fault Diagnosis, *Proceeding of the 2016 International Symposium on Flexible Automation*, 2016, pp. 390-396.
- [9] Fontes, C.H. and Pereira, O., Pattern Recognition in Multivariate Time Series – A Case Study Applied to Fault Detection in a Gas Turbine, *Engineering Applications of Artificial Intelligence*, 2016, Vol. 49, pp. 10-18.
- [10] Fu, T.C., A Review on Time Series Data Mining, *Engineering Applications of Artificial Intelligence*, 2011, Vol. 24, No. 1, pp. 164-181.
- [11] García, S., Luengo, J., Sáez, J.A., López, V., and Herrera, F., A Survey of Discretization Techniques: Taxonomy and Empirical Analysis in Supervised Learning, *IEEE*

- Transactions on Knowledge and Data Engineering*, 2013, Vol. 25, No. 4, pp. 734-750.
- [12] Ghosh, A., Wang, G.-N., and Lee, J., A Novel Automata and Neural Network Based Fault Diagnosis System for PLC Controlled Manufacturing Systems, *Computers & Industrial Engineering*, 2020, Vol. 139, pp. 1-16.
- [13] Hodge, V. and Austin, J., A Survey of Outlier Detection Methodologies, *Artificial Intelligence Review*, 2014, Vol. 22, No. 2, pp. 85-126.
- [14] Jung, H. and Kim, J.-W., A Machine Learning Approach for Mechanical Motor Fault Diagnosis, *J. Soc. Korea Ind. Syst. Eng.*, 2017, Vol. 40, No. 1, pp. 57-64.
- [15] Keogh, E., Chu, S., Hart, D., and Pazzani, M., An Online Algorithm for Segmenting Time Series, *Proceedings of the 2001 IEEE International Conference on Data Mining*, 2001, San Jose, CA, U.S., pp. 289-296.
- [16] Keogh, E. and Kasetty, S., On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration, *Data Mining and Knowledge Discovery*, 2003, Vol. 7, No. 4, pp. 349-371.
- [17] Kim, G.E., Hwang, H.S., Lee, K.K., and Han, S.H., Fault Detection of Rolling Bearing Using Performance Index of Statistical Process Control, *The Transactions of Korean Society of Mechanical Engineering A*, 2021, Vol. 45, No. 1, pp.51-57.
- [18] Kim, H., Lee, J., and Kim, J., Development of State Diagnosis Algorithms Using Statistical Method for Online Sensors in A2/O Process, *Journal of Korean Society of Environmental Engineers*, Vol. 41, No. 11, pp. 672-636.
- [19] Kim, J.W., Lee, J.H., Seok, H., Choi, S.D., and Hur, J.W., A Study of the Prognostics of Inverter Failure Using RF and KNN, *Proceedings of Spring Conference of The Korean Society of Manufacturing Process Engineers*, 2020, p. 96,
- [20] Lee, G.H., Shin, B.C., and Hur, J.W., Fault Classification of Gear Pumps Using SVM, *Journal of Applied Reliability*, 2020, Vol. 20, No. 2, pp. 187-196.
- [21] Lee, I., Diagnostic System Development for State Monitoring of Induction Motor and Oil Level in Press Process System, *Journal of Korean Institute of Intelligent Systems*, 2009, Vol. 19, No. 5, pp. 706-712.
- [22] Lim, J.-K. and Yoon, H.-J., A Study on the Feature Extraction for the Rolling Stock Door Using the Current Value of The Motor and the Selection of a Failure Diagnosis Prediction Algorithm, *The Transactions of the Korean Institute of Electrical Engineers*, 2021, Vol. 21, No. 1, pp. 96-101.
- [23] Liu, H., Hussain, F., Tan, C.L., and Dash, M., Discretization: An Enabling Technique, *Data Mining and Knowledge Discovery*, 2002, Vol. 6, No. 4, pp. 393-423.
- [24] Luo, R.C., Sensor Technologies and Microsensor Issues for Mechatronics Systems, *IEEE/ASME Transactions on Mechatronics*, 1996, Vol. 1, No. 1, pp. 39-49.
- [25] Namgung, K., Yoon, S., Baek, S., and Kim, D.-Y., Estimating System State through Similarity Analysis of Signal Patterns, *Sensors*, 2020, Vol. 20, No. 23, pp. 1-13.
- [26] Phaladiganon, P., Kim, S.B., Chen, V.C. , and Jiang, W. Principal Component Analysis-based Control Charts for Multivariate Nonnormal Distribution, *Expert Systems with Applications*, 2013, Vol. 40, pp. 3044-3054.
- [27] Rahman, M.A., Rashid, M.A., and Ahmad, M., Selecting the Optimal Conditions of Savitzky-Golay Filter for fNIRS Signal, *Biocybenetics and Biomedical Engineering*, 2019, Vol. 39, No. 3, pp. 624-637.
- [28] Saxena, A. and Goebel, K., Turbofan Engine Degradation Simulation Data Set, *NASA Ames Prognostics Data Repository* (<http://ti.arc.nasa.gov/project/prognostic-data-repository>), 2008, NASA Ames Research Center, Moffett Field, CA.
- [29] Saxena, A., Goebel, K., Simon, D., and Eklund, N., Damage Propagation Modeling for Aircraft Engine Run-to-failure Simulation, *Proceedings of the International Conference on Prognostics and Health Management*, 2008, Denver, CO, U.S., pp. 1-9.
- [30] Shewhart, M., Application of Machine Learning and Expert Systems to Statistical Process Control (SPC) Chart Interpretation, *Proceedings of the 2nd C Language Integrated Production System Conference*, 1991, Houston, TX, U.S. pp. 123-138.
- [31] Tamboli, K., Sheth, S., Shah, V., Gandhi, C., Amin, N., and Modi, V., Design and Development of a Mechatronic System for the Measurement of Railway Tracks, *The Industrial Daily Journal*, 2015, Vol. 43, No. 200, pp. 174-180.
- [32] Wilson, S.J., Data Representation for Time Series Data Mining: Time Domain Approaches, *Wiley Interdisciplinary Reviews: Computational Statistics*, 2017, Vol. 9, No. 1, pp. e1392:1-6.

- [33] Xu, Y. and Liu, J., High-speed Train Fault Detection with Unsupervised Causality-Based Feature Extraction Methods, *Advanced Engineering Informatics*, 2021, Vol. 49, p. 101312.
- [34] Yu, G.S. and Moon, Y.M., Development trend of smart factory facility diagnosis and predictive maintenance technology, *The Journal of The Korean Institute of Communication Sciences*, 2020, Vol. 37, No. 7, pp. 36-42.

ORCID

Sujeong Baek | <http://orcid.org/0000-0002-9387-6217>